METplus Verification and Diagnostics Framework for Model Evaluation Across Scales

Tara Jensen[1,2], John Opatz[1,2], Christina Kalb[1,2], Daniel Adriaansen[1,2], Kathryn Newman[1,2], Michelle Harrold[1,2], Mrinal Biswas[1,2], Tracy Hertneky[1,2], Will Mayfield[1,2], Weiwei Li[1,2], Brianne Nelson[1,2], Jonathan Vigh[1,2], Molly Smith[3,4], Jason English[3,4], Louisa Nance[1,2], Barbara Brown[1], Michael Ek[1,2]

[1]National Center for Atmospheric Research, [2]Developmental Testbed Center, [3]Cooperative Institute for Research in Environmental Sciences, [4]NOAA Global Systems Laboratory

Corresponding Author: jensen@ucar.edu

Introduction

The METplus system is a suite of verification and diagnostic tools in a consistent framework that are designed to facilitate consistent computation of statistics and metrics across applications and institutions. The highly configurable Python wrappers provide low-level workflow around the core Model Evaluation Tools (MET) package for computing verification statistics. Additional components of METplus include an Analysis Suite made up of a data input-output library (METdataio), an aggregation and synthesis tool (METcalcpy), and plotting (METplotpy). The tools are more fully discussed in Brown et. al (2021) and provide great flexibility to evaluate a range of numerical prediction across temporal scales, spatial scales, and model applications. To address the needs of a cadre of international research laboratories and operational centers, METplus is also being enhanced to provide systematic evaluation and diagnostics of many of the coupled components within an Earth system modeling framework.

Temporal and Spatial Scales

METplus was designed to work like most other Linux/Unix based tools, with each component being focused on a small subset of capability to allow for maximum flexibility in setting up the tools. Unless specifically designed to accept a time series of data, the tools focus on computing statistics for a give valid time and allow for aggregation of statistics, diagnostic attributes, or metrics, over the appropriate temporal scales for a model application. This allows the same tools to be used for evaluation of short-range (1-10 minute, hourly, daily), medium-range (3-14 days), sub-seasonal (weeks 2-4), seasonal (up to 9 months), and climate (yearly to multi-decadal) simulations. The capability includes the use of appropriate user-defined climatologies, thresholds, masking regions to define areas of interest, and interpolation methods.

Applications

The applications METplus has been applied to vary from evaluation forecasts for renewable energy, fire weather, severe weather, extremes, marine (standard variables as well as phenomena like chlorophyll), cryosphere, extra-tropical and tropical cyclones, monsoons, droughts, clouds, dust, aerosols, air quality fields, satellite brightness temperatures, land modeling diagnostics, and ionospheric fields for space weather. The tools are used in both research and operations. Current operational partners include the National Oceanic and Atmospheric Administration (NOAA) National Center for Environmental Prediction (NCEP) Centers, the United States Air Force, the Met Office for the United Kingdom, the Australian Bureau of Meteorology, the Indian National Center for Medium Range Weather Forecasting, and many of the other associated Unified Model partnership. Additionally, METplus is being adopted by US Naval Research Laboratory and US Army Research Laboratory, and the United Arab Emirates National Center for Meteorology.

Current and Future Capability

In METplus v5.1, released June 2023, there is support for over 150 traditional statistics and diagnostics methods. Table 1 provides a synopsis of what is included in the METplus framework. Future work is focused on optimizing the use of memory, adding support for parallelized computing and unstructured grids, developing database applications to store large quantities of forecast-observation matched pairs for analysis, and including more diagnostics for non-atmospheric components of an Earth system model.

Traditional	
Grid-Stat, Point-Stat, Series-Analysis Contingency table (CTS), Continuous, Probability forecast statistics, SEEPS	**Ensemble-Stat** CRPS, CRPSS, Rank probability, Prob. Integral Transform, and Relative position histograms, Spread/Skill, Ignorance
Spatial	
MODE Location and Geometric attribute differences, Intersection area, Intensity distributions, CTS measures for objects, *****Available for single and multi-variate fields***	**MODE-Time Domain** Time and location differences, Volume and Velocity differences, Intersection volume, Intensity distributions and differences
Wavelet-Stat Mean Square Error by scale, Energy by scale, Intensity-scale skill score	**Grid-Stat and Point-Stat** Fraction Skill Score, High Resolution Analysis, Distance Measures, Mean Error Distance, Baddeley, Hausdorff, Zhu, Fourier Decomposition
Diagnostics	
Grid-Diag Distribution of fields to assess multivariate relationships between fields	**Feature Relative** Used to assess presence of systematic errors associated with features or events
Physics Tendencies Computation of Physics tendencies (vertical cross section and plan-view)	
Tropical Cyclones Application	
TC-Pairs, TC-Stat, TC-Dland Track error (along, cross, total), Intensity (pressure, wind), Rapid Intensification/weakening errors, CTS measures for TC genesis	**TC-Gen** CTS measures for TC genesis, Spatial representation of TC genesis density function, TC density function
TC-RMW Errors and diagnostics in Radius of maximum wind projection	**TC-Diag** Errors and diagnostics in
S2S Application	
Realtime Multivariate Madden Julien Oscillation (MJO) Index (RMM) 1, RMM2, Outgoing Longwave Radiation (OLR) MJO Index, MJO-El Nino Southern Oscillation (ENSO) Index	
Identification of Weather Regimes and Blocking Regimes, Hovmoeller Diagrams, Zonal and Meridional Means, Empirical Orthogonal Functions (EOFs), Space-Time Coherence (or Cross-Spectra) Plots	
Statistical Synthesis Tools	
Scorecards, Contour Plots	Performance Diagrams, Taylor Diagrams

Table 1. High level summary of traditional statistics, spatial methods and model diagnostics, and Tropical Cyclone application support. MET Tool names are in blue with capability listed in black.

Summary

METplus is a state-of-the-science community verification and diagnostic package that is used by over 3,000 US and international institutions spanning the public, private, and academic sectors. It provides the Earth system modeling research community with the ability to share, and hence, standardize evaluation across entities. More information can be found at: dtcenter.org/community-code/metplus

References

Brown, B., T. G. Jensen, et al., **2021**: The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bull. Amer. Meteor. Soc.*, E782-E807. DOI: https://doi-org.cuucar.idm.oclc.org/10.1175/BAMS-D-19-0093.1